Optimal Family of q-ary Codes Obtained From a Substructure of Generalised Hadamard Matrices

Carl Bracken, Yeow Meng Chee and Punarbasu Purkayastha
Coding and Cryptography Research Group
Division of Mathematical Sciences,
School of Physical and Mathematical Sciences,
Nanyang Technological University,
Singapore

***Abstract*—In this article we construct an infinite family of linear error correcting codes over $\mathbb{F}_q$ for any prime power q. The code parameters are**

$$[q^{2t}+q^{t-1}-q^{2t-1}-q^t, 2t+1, q^{2t}+q^{2t-2}+q^{t-1}-2q^{2t-1}-q^t]_q,$$

for any positive integer t. This family is a generalisation of the optimal self-complementary binary codes with parameters

$$[2u^2-u, 2t+1, u^2-u]_2,$$

where $u = 2^{t-1}$. The codes are obtained by considering a submatrix of a specially constructed generalised Hadamard matrix. The optimality of the family is confirmed by using a recently derived generalisation of the Grey-Rankin bound when $t > 1$, and the Griesmer bound when $t = 1$.

***Index Terms*—Generalised Hadamard matrix, Grey-Rankin bound.**

I. Introduction

We use the usual notation $(n, M, d)_q$ to denote an error correcting code in $\mathbb{F}_q^n$ of size M and minimum distance d. If the code is a linear subspace of dimension k, then we denote it as $[n, k, d]_q$. As this work is a generalisation of a family of binary codes, we begin by reviewing the binary case.

Definition 1: A Hadamard matrix H is an n by n matrix with entries in $\{1, -1\}$ such that,

$$HH^T = nI,$$

where I denotes the n by n identity matrix.

It can be easily demonstrated that for a Hadamard matrix to exist n must be a multiple of 4. It is conjectured, with strong evidence, that a Hadamard matrix exists for all n divisible by 4.

Definition 2: A binary error correcting code C is said to be *self-complementary* if for all words $x \in C$ we have $x+\mathbf{1} \in C$, where $\mathbf{1}$ is the all-1 vector $(1, \ldots, 1)$.

By changing the symbols in the rows of H from 1 and -1 to 0 and 1, then adding to this set of rows the complements of the rows and puncturing this code in one coordinate (by deleting any column) we obtain an optimal self-complementary code with parameters $(n-1, 2n, \frac{n}{2}-1)_2$. This is known as a Hadamard code. The optimality of a self-complementary code can be tested with the Grey-Rankin bound which states that

$$M \leq \frac{8d(n-d)}{n-(n-2d)^2},$$

provided the RHS of the inequality is positive. The Hadamard code meets this bound with equality. A result of McGuire [10] states that a self-complementary code meeting this bound must be either a Hadamard code or must arise from a quasi-symmetric design with specified intersection numbers. We do not discuss the details of the quasi-symmetric designs here, but we are interested in the codes they give. In [5] and [8] there are constructions of these designs, which yield codes with parameters

$$(2u^2-u, 8u^2, u^2-u)_2,$$

for $u = 2^m$, m an integer, $m > 1$. These codes are believed to exist for all even u. They have been shown to exist whenever there exists a u times u Hadamard matrix (see [2] and [3]). Also, there is a construction for $u = 6$ [4] but all other cases where u is equivalent to 2 modulo 4 are open. These constructions use the structure of Hadamard matrices to obtain the quasi-symmetric designs and hence the codes. When $u = 2^{t-1}$, $t > 1$ and the codes have minimum 2-rank, they give the linear family of codes

$$[2u^2-u, 2t+1, u^2-u]_2.$$

This is the family of codes that we are going to generalise to the q-ary Hamming space by using the structure of generalised Hadamard matrices.

II. Preliminaries

Definition 3: Let G be a group of order g and let λ be a positive integer. A *generalised Hadamard matrix* $GH(g, \lambda)$ over the group G is a $\lambda g \times \lambda g$ matrix such that the pairwise difference of any two rows of the matrix contains every element of G exactly λ times.

For more about generalised Hadamard matrices we refer the reader to [6] .

The generalised Hadamard matrix is called normalised if the first row and the first column consist of only the identity element of G. In this work we consider G to be the additive group in the finite field $\mathbb{F}_{q^t}$ for $t \geq 1$, where q is the power of a prime. The generalised Hadamard matrix is also considered as a code where each row is a codeword of the code. For a $GH(g, \lambda)$, we get a code of length λg, size λg, and distance $\lambda g - \lambda$.

There has been more than one generalisation of the Grey Rankin bound. In this article we are only using the most recent one from Bassalygo, et. al. [1]. This generalised Grey-Rankin bound is stated as follows.

Theorem 2.1: [1] Let $\mathcal{C}$ be an $(n, M, d)_q$ code such that it can be partitioned into trivial maximal subcodes $(n, q, n)_q$. Then the size of the code satisfies

$$M \leq \frac{q^2(n-d)(qd-(q-2)n)}{n-((q-1)n-qd)^2}.$$

The condition that a code can be partitioned into trivial maximal subcodes $(n, q, n)_q$ is equivalent to the property, for all words $x \in C$ we have $x + \mathbf{1} \in C$. So this can be thought of as a generalisation of the self-complementary property of binary codes. Any linear code that contains the all one vector also has this property. We also note, when $q = 2$ this bound reduces to the binary Grey-Rankin bound.

For an $[n, k, d]_q$ linear code, the Griesmer bound gives the length $n(k, d)$ of the shortest linear code with dimension k and minimum distance d.

Theorem 2.2: [11] Let $\mathcal{C}$ be an $[n, k, d]_q$ linear code. Then

$$n(k,d) \geq d + n\Big(k-1, \left\lceil \frac{d}{q} \right\rceil\Big) = d + \left\lceil \frac{d}{q} \right\rceil + \cdots + \left\lceil \frac{d}{q^{k-1}} \right\rceil.$$

III. Construction of $GH(q^k, q^{2t-k})$

In this section we show how to construct a specific Generalised Hadamard matrix with parameters $GH(q^k, q^{2t-k})$ from a $GH(q^t, 1)$. The construction is achieved by considering the $GH(q^t, 1)$ as a code and then performing code operations such as extension and concatenation.

Let H denote the normalised generalised Hadamard matrix $GH(q^t, 1)$ with entries

$$H(i,j) = \begin{cases} 0, & i = 0, \text{ or } j = 0, \\ \alpha^{i+j}, & 1 \leq i, j \leq q^t - 1, \end{cases}$$

where α is a primitive element of $\mathbb{F}_{q^t}$. This matrix, when considered as a codematrix, is a linear code with parameters $[q^t, t, q^t - 1]_{q^t}$. The linearity of H can be proved as follows. Let $r_i = (0, \alpha^i, \alpha^{i+1}, \ldots, \alpha^{i+q^t-2})$ denote a row of the matrix H. We let $r_0 = \mathbf{0}$ denote the first row. Then observe that

$$\begin{aligned} r_i + \alpha^l r_j &= (\alpha^i + \alpha^l \alpha^j)(0, 1, \alpha, \ldots, \alpha^{q^t-2}) \\ &= (0, \alpha^k, \alpha^{k+1}, \ldots, \alpha^{k+q^t-2}) \\ &= r_k, \end{aligned}$$

where $\alpha^k = \alpha^i + \alpha^l \alpha^j$.

The fact that H is a generalised Hadamard matrix follows from its linearity. Consider the code $\mathcal{C}_H$ obtained by taking H along with all its cosets $H + \beta\mathbf{1}$ and $\beta \in \mathbb{F}_{q^t}$. The code $\mathcal{C}_H$ is also linear and has parameters $[q^t, 2t, q^t - 1]_{q^t}$. This code is optimal because it satisfies the generalised Grey-Rankin bound with equality [1]. Extend the code $\mathcal{C}_H$ by appending an extra element to every element of $\mathcal{C}_H$ as described below. For each $i = 1, \ldots, q^t - 1$, the row r_i of H is extended by appending the element α^i and the first row of H is extended by appending the element 0. Similarly the row in the coset $H + \beta\mathbf{1}$ which is obtained from row r_i of H, is extended by appending α^i to it, and the row in the coset $H + \beta\mathbf{1}$ corresponding to r_0 of H is extended by appending 0 to it. Denote this extended code by $\mathcal{C}_H^+$. The following lemma gives the parameters of $\mathcal{C}_H^+$.

Lemma 3.1: The code $\mathcal{C}_H^+$ is an optimal nonlinear equidistant code with parameters $(q^t + 1, q^{2t}, q^t)_{q^t}$.

Proof: The extended code clearly has length $q^t + 1$ and size q^{2t}. To determine the distances in the extended code we first determine the distances in the code $\mathcal{C}_H$. The distance between elements $r_i + \beta\mathbf{1}, r_j + \beta'\mathbf{1}$ of cosets $H + \beta\mathbf{1}$ and $H + \beta'\mathbf{1}$, $\beta, \beta' \in \mathbb{F}_{q^t}$, respectively, in $\mathcal{C}_H$ is given by

$$d(r_i + \beta\mathbf{1}, r_j + \beta'\mathbf{1}) = \begin{cases} q^t - 1, & \text{if } \beta = \beta', i \neq j \\ q^t - 1, & \text{if } \beta \neq \beta', i \neq j \\ q^t, & \text{if } \beta \neq \beta', i = j. \end{cases}$$

Thus, the distance between elements $r_i + \beta\mathbf{1}$ and $r_j + \beta'\mathbf{1}$ of the extended code $\mathcal{C}_H^+$ is given by

$$d(r_i + \beta\mathbf{1}, r_j + \beta'\mathbf{1}) = \begin{cases} q^t, & \text{if } \beta = \beta', i \neq j \\ q^t, & \text{if } \beta \neq \beta', i \neq j \\ q^t, & \text{if } \beta \neq \beta', i = j. \end{cases}$$

Hence, the code is equidistant. The code $\mathcal{C}_H^+$ is optimal because it satisfies the Plotkin bound with equality. A code with parameters $(n, M, d)_q$ satisfies the Plotkin bound if

$$M \leq \left\lfloor \frac{d}{d - (q-1)n/q} \right\rfloor.$$

Straightforward calculations show that the RHS is exactly q^{2t} for the code $\mathcal{C}_H^+$. ■

Next, we construct a family of generalised Hadamard matrices $GH(q^k, q^{2t-k})$, for $k = 1, \ldots, t$ from the matrix $H = GH(q^t, 1)$ and from the code $\mathcal{C}_H^+$. This is obtained by concatenating the code $\mathcal{C}_H^+$ with a punctured matrix obtained from H.

Consider a linear projection of each element (in the additive group of $\mathbb{F}_{q^t}$) of H on to the additive subgroup in $\mathbb{F}_{q^k}$ for any $k \in \{1, \ldots, t\}$. This projection can be achieved by first considering the field elements as vectors, with respect to a fixed basis, and then mapping the last $t - k$ coordinates to zero. This gives a generalised Hadamard matrix $H(k) = GH(q^k, q^{t-k})$ [6].

We construct a matrix $H^-(k)$ from $H(k)$ by removing its first column. When considered as a code $H^-(k)$ has parameters $[q^t-1, t, q^t-q^{t-k}]_{q^k}$. We concatenate $H^-(k)$ with $\mathcal{C}_H^+$ by replacing every element α^i in $\mathcal{C}_H^+$ by the row r_i of $H^-(k)$, $i=1,\dots,q^t-1$, and by replacing the element 0 in $\mathcal{C}_H^+$ by the first row of $H^-(k)$. Denote the concatenated code by $H^-(k)\circ\mathcal{C}_H^+$. Finally, extend the concatenated codematrix $H^-(k)\circ\mathcal{C}_H^+$ by prepending an all-zero column to obtain the codematrix $H^2(k)$.

Proposition 3.2: The matrix $H^2(k)$ is a generalised Hadamard matrix $GH(q^k, q^{2t-k})$.

Proof: Since the distance between any two rows in $\mathcal{C}_H^+$ is q^t, the two rows are equal in exactly one position. Thus, the number of zeroes in the difference of two rows of $\mathcal{C}_H^+\circ H^-(k)$ is exactly $q^t(q^{t-k}-1)+q^t-1$. The extended code contributes one extra zero to this count.

For any nonzero element, the number of such non-zero elements in the difference of two rows of $H^-(k)\circ\mathcal{C}_H^+$ is exactly $q^t(q^{t-k})$ since $H^-(k)$ contains all the nonzero elements equally often. This proves that the code $H^2(k)$ is a generalised Hadamard matrix. ∎

IV. Construction of optimal linear codes

In this section we describe the construction of a family of optimal linear codes with parameters

$$[d(q^t-1), 2t+1, d(d-1)]_q,$$

where $d=q^t-q^{t-1}$. This optimal linear code is obtained from $H^2=H^2(1)$. The construction is performed by first concatenating a punctured code obtained from $\mathcal{C}_H$ with the code $H^-(1)$. Then the resulting concatenated code is augmented by the all-one vector.

Fix a positive integer $s\geq d$. Puncture the code $\mathcal{C}_H$ such that the resulting code contains only the first s coordinates (more generally, one may puncture it on any set of coordinates such that the resulting code has exactly s coordinates). Denote the punctured code by $\mathcal{C}_H^*$. It has parameters $[s, 2t, s-1]_{q^t}$.

Consider the concatenation of the code $\mathcal{C}_H^*$ with the code $H^-(1)$. We denote the concatenated code by $H^-(1)\circ\mathcal{C}_H^*$. Since the codes $H^-(1)$ and $\mathcal{C}_H^*$ are $\mathbb{F}_q$-linear, the concatenated code $H^-(1)\circ\mathcal{C}_H^*$ is also a linear code over $\mathbb{F}_q$ (see [7]). Consider the code $\mathcal{C}(s)$ obtained by augmenting the concatenated code by the all-one vector $\mathbf{1}$ of length $s(q^t-1)$. Note that the code $\mathcal{C}(s)$ can also be obtained by puncturing the matrix $H^2(1)$ appropriately, and then augmenting it by the all-one vector $\mathbf{1}\in\mathbb{F}_q^{s(q^t-1)}$. The proposition below establishes the parameters of the code $\mathcal{C}(s)$.

Lemma 4.1: The code $\mathcal{C}(s)$ has parameters

$$[s(q^t-1), 2t+1, (d-1)s]_q,$$

where $d=q^t-q^{t-1}$. The code $\mathcal{C}(s)$ has five distances (in increasing order):
$(d-1)s, d(s-1), sd+q^t-s-d, sd$, and $s(q^t-1)$.

Proof: The distances in the code $\mathcal{C}(s)$ are derived from the distances in the component codes $H^-(1)$ and $\mathcal{C}_H$. Any two rows of $H^-(1)$ have distance $d=q^t-q^{t-1}$. Two rows of $\mathcal{C}_H^*$ have distance either s or $s-1$. Thus two rows of $H^-(1)\circ\mathcal{C}_H^*$ have distance either sd or $(s-1)d$. Without loss of generality let $\mathbf{r}_0$ be a row from $H^-(1)\circ\mathcal{C}_H^*$ and $\mathbf{r}_1$ be a row from the set of codewords $\{\mathbf{c}+\mathbf{1} : \mathbf{c}\in H^-(1)\circ\mathcal{C}_H^*\}$. If $\mathbf{r}_1=\mathbf{r}_0+\mathbf{1}$ then the distance between them is $s(q^t-1)$. Otherwise, we can write $\mathbf{r}_1=\mathbf{r}'+\mathbf{1}$, where $\mathbf{r}'\neq\mathbf{r}_0$. Then the distance between the codewords is given as

$$d(\mathbf{r}_0,\mathbf{r}_1)=\begin{cases} s(q^t-1), & \text{if } \mathbf{r}_1=\mathbf{r}_0+\mathbf{1},\\ s(d-1), & \text{if } d(\mathbf{r}_0,\mathbf{r}')=s,\\ sd+n-s-d, & \text{if } d(\mathbf{r}_0,\mathbf{r}')=s-1,\end{cases}$$

where the last case follows because the distance between the vectors is $(s-1)(d-1)+n-1$. This establishes the distances in the concatenated code $\mathcal{C}(s)$. The minimum distance of the code can be inferred from the inequality $s(d-1)\leq(s-1)d$, for $s\geq d$. ∎

The optimal linear code is now obtained by reducing the number of distances in the code to four distances. This is effected by choosing $s=d$. It now has parameters

$$[q^{2t}+q^{t-1}-q^{2t-1}-q^t, 2t+1, q^{2t}+q^{2t-2}+q^{t-1}-2q^{2t-1}-q^t]_q,$$

Denote this linear code by $\mathcal{C}=\mathcal{C}(d)$. Below we prove the optimality of the code by comparing it to the generalised Grey-Rankin bound for $t>1$, and to the Griesmer bound for $t=1$. When $t=1$ the code has parameters;

$$[q^2-2q+1, 3, q^2-3q+2]_q.$$

Theorem 4.2: For a linear code containing the all one vector, the code $\mathcal{C}=\mathcal{C}(d)$ is optimal when $t>1$. When $t=1$, $\mathcal{C}(d)$ is an optimal linear code.

Proof: Let $n=q^t$ and $d=n-n/q$. Let C be any linear code with parameters $[N,K,D]_q$ where $N=d(n-1), D=d(d-1)$, and K is the dimension of the code. Let M be the size of C.

We first prove the result for $t>1$. Substituting the values of N,D from the parameters of C gives us the generalised Grey-Rankin upper bound

$$\begin{aligned} M &\leq \frac{q^2\big(d(n-1)-d(d-1)\big)\big(qd(d-1)-(q-2)d(n-1)\big)}{d(n-1)-\big((q-1)d(n-1)-qd(d-1)\big)^2}\\ &= q^2\frac{dn(n-2)/q}{n/q-1}\\ &= qn^2\left(q-\left(1-\frac{(q-2)(q-1)}{n-q}\right)\right). \end{aligned}$$

We have used the fact that $d=n-n/q$ in arriving at the above expression. Also note that $n=q^t$ and hence the term multiplying $qn^2=q^{2t+1}$ is strictly less than q. To show this, substitute $n=q^t$ and note that for $t\geq 2$,

$$\frac{(q-2)}{q}\frac{(q-1)}{q^{t-1}-1}<1.$$

For the linear code $\mathcal{C}$ with dimension $K = 2t+1$, it is optimal if no larger linear code with dimension $2t+2$ can be found. The generalised Grey-Rankin bound thus proves the optimality of $\mathcal{C}$.

For $t = 1$, we use the Griesmer bound and show that $\mathcal{C}$ satisfies the Griesmer bound with equality. Note that $K = 2t+1 = 3$. To satisfy the Griesmer bound with equality we need to show that

$$N = D + \left\lceil \frac{D}{q} \right\rceil + \left\lceil \frac{D}{q^2} \right\rceil .$$

With $D = (q-1)(q-2)$ and $N = (q-1)^2$ it is readily verified that the RHS of the above equation is $(q-1)(q-2)+(q-2)+1$ which equals the LHS. ■

References

[1] L. Bassalygo, S. Dodunekovt, T. Hellesetht and V. Zinoviev, *On a new q-ary combinatorial analog of the binary Grey-Rankin bound and codes meeting this bound*, IEEE Information Theory Workshop, Punta del Este, Uruguay, March 13–17, 2006, 278–282.

[2] C. Bracken, *New classes of self-complementary codes and quasi-symmetric designs* , Designs, Codes and Cryptography, 2006, (**41**), 319–323.

[3] C. Bracken, *Pseudo Quasi-3 Designs and their Applications to Coding Theory*, Journal of Combinatorial Designs, **17**, 2009, 411-418.

[4] C. Bracken, G. McGuire, H. N. Ward, *New quasi-symmetric designs constructed using mutually orthogonal Latin squares and Hadamard matrices*, Designs, Codes and Cryptography, 2006, **41**, 195-198.

[5] P. J. Cameron, *Near-regularity conditions for designs*, Geom. Ded., 1973, **2**, 213-223.

[6] C. J. Colbourn and J. H. Dinitz, editors. *Handbook of Combinatorial Designs (second edition)*, Chapman & Hall/CRC, Boca Raton, FL, 2007, 301-306.

[7] I. I. Dumer, *Concatenated codes and their multilevel generalizations*, Handbook of Coding Theory, Vol. II, V. S. Pless, W. C. Huffman and R. A. Brualdi eds., North-Holland, Amsterdam, 1998, 1911–1988.

[8] D. Jungnickel and V. D. Tonchev, *Exponential number of quasi-symmetric SDP designs and codes meeting the Grey-Rankin bound*, Designs, Codes and Cryptography, 1991, (**1**), 247-253.

[9] C. Mackenzie and J. Seberry, *Maximal q-ary codes and Plotkin bound*, Ars Combinatorica, 1988, (**26B**), 37–50.

[10] G. McGuire, *Quasi-symmetric designs and codes meeting the Grey-Rankin bound*, J. Combin. Theory Ser. A, 1997, **72**.

[11] G. Solomon and J. J. Stiffler, *Algebraically punctured cyclic codes*, Information and Control, 1965, (**8**), no. 2, 170-179.